

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/790,043ADATE: 10/22/97
TIME: 12:56:35

INPUT SET: S21089.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

#7/7R
07/04/98

SEQUENCE LISTING

ENTERED

(1) General Information

(i) APPLICANT: Lonsdale, John
Milner, Peter
Payne, David
Pearson, Stewart

(ii) TITLE OF THE INVENTION: Novel FabI

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SmithKline Beecham Corporation
(B) STREET: 709 Swedeland Road
(C) CITY: King of Prussia
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19406-0939

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 28-August-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/024845
(B) FILING DATE: 28-AUG-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Gimmi, Edward R
(B) REGISTRATION NUMBER: 38,891
(C) REFERENCE/DOCKET NUMBER: GM50005

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 610-270-4478
(B) TELEFAX: 610-270-5090
(C) TELEX:

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/790,043ADATE: 10/22/97
TIME: 12:56:37

INPUT SET: S21089.raw

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 256 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Leu	Asn	Leu	Glu	Asn	Lys	Thr	Tyr	Val	Ile	Met	Gly	Ile	Ala	Asn
1				5					10					15	
Lys	Arg	Ser	Ile	Ala	Phe	Gly	Val	Ala	Lys	Val	Leu	Asp	Gln	Leu	Gly
			20					25					30		
Ala	Lys	Leu	Val	Phe	Thr	Tyr	Arg	Lys	Glu	Arg	Ser	Arg	Lys	Glu	Leu
			35					40					45		
Glu	Lys	Leu	Leu	Glu	Gln	Leu	Asn	Gln	Pro	Glu	Ala	His	Leu	Tyr	Gln
			50					55					60		
Ile	Asp	Val	Gln	Ser	Asp	Glu	Glu	Val	Ile	Asn	Gly	Phe	Glu	Gln	Ile
					70					75					80
Gly	Lys	Asp	Val	Gly	Asn	Ile	Asp	Gly	Val	Tyr	His	Ser	Ile	Ala	Phe
				85					90					95	
Ala	Asn	Met	Glu	Asp	Leu	Arg	Gly	Arg	Phe	Ser	Glu	Thr	Ser	Arg	Glu
			100					105					110		
Gly	Phe	Leu	Leu	Ala	Gln	Asp	Ile	Ser	Ser	Tyr	Ser	Leu	Thr	Ile	Val
			115					120					125		
Ala	His	Glu	Ala	Lys	Lys	Leu	Met	Pro	Glu	Gly	Gly	Ser	Ile	Val	Ala
			130					135					140		
Thr	Thr	Tyr	Leu	Gly	Gly	Glu	Phe	Ala	Val	Gln	Asn	Tyr	Asn	Val	Met
						150				155					160
Gly	Val	Ala	Lys	Ala	Ser	Leu	Glu	Ala	Asn	Val	Lys	Tyr	Leu	Ala	Leu
				165					170					175	
Asp	Leu	Gly	Pro	Asp	Asn	Ile	Arg	Val	Asn	Ala	Ile	Ser	Ala	Gly	Pro
				180				185					190		
Ile	Arg	Thr	Leu	Ser	Ala	Lys	Gly	Val	Gly	Gly	Phe	Asn	Thr	Ile	Leu
			195					200					205		
Lys	Glu	Ile	Glu	Glu	Arg	Ala	Pro	Leu	Lys	Arg	Asn	Val	Asp	Gln	Val
			210					215					220		
Glu	Val	Gly	Lys	Thr	Ala	Tyr	Leu	Leu	Ser	Asp	Leu	Ser	Ser	Gly	
					230					235				240	
Val	Thr	Gly	Glu	Asn	Ile	His	Val	Asp	Ser	Gly	Phe	His	Ala	Ile	Lys
				245					250					255	

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 771 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/790,043ADATE: 10/22/97
TIME: 12:56:39

INPUT SET: S21089.raw

100 (ii) MOLECULE TYPE: Genomic DNA

101

102 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

103

104	ATGTTAAATC	TTGAAAACAA	AACATATGTC	ATCATGGGAA	TCGCTAATAA	GCGTAGTATT	60
105	GCTTTTGGTG	TCGCTAAAGT	TTTAGATCAA	TTAGGTGCTA	AATTAGTATT	TACTTACCGT	120
106	AAAGAACGTA	GCCGTAAAGA	GCTTGAAAAA	TTATTAGAAC	AATTAAATCA	ACCAGAAGCG	180
107	CACTTATATC	AAATTGATGT	TCAAAGCGAT	GAAGAGGTTA	TTAATGGTTT	TGAGCAAATT	240
108	GGTAAAGATG	TTGGCAATAT	TGATGGTGTA	TATCATTCAA	TCGCATTTGC	TAATATGGAA	300
109	GACTTACGCG	GACGCTTTTC	TGAAACTTCA	CGTGAAGGCT	TCTTGTTAGC	TCAAGACATT	360
110	AGTTCTTACT	CATTAACAAT	TGTGGCTCAT	GAAGCTAAAA	AATTAATGCC	AGAAGGTGGT	420
111	AGCATTTGTTG	CAACAACATA	TTTAGGTGGC	GAATTCGCAG	TTCAAAATTA	TAATGTGATG	480
112	GGTGTGCTA	AAGCGAGCTT	AGAAGCAAAT	GTTAAATATT	TAGCATTAGA	CTTAGGTCCCT	540
113	GATAATATTC	GCGTTAATGC	AATTTTCAGCT	GGTCCAATCC	GTACATTAAG	TGCAAAAGGT	600
114	GTGGGTGGTT	TCAATACAAT	TCTTAAAGAA	ATCGAAGAGC	GTGCACCTTT	AAAACGTAAC	660
115	GTTGATCAAG	TAGAAGTAGG	TAAAACAGCG	GCTTACTTCT	TAAGTGACTT	ATCAAGTGCG	720
116	GTTACAGGTG	AAAATATTCA	TGTAGATAGC	GGATTCCACG	CAATTAAATA	A	771

117

118

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/790,043A

DATE: 10/22/97
TIME: 12:56:40

INPUT SET: S21089.raw

Line

Error

Original Text